

Relatedness of the Incidence Decay with Exponential Adjustment (IDEA) Model, “Farr’s Law” and Compartmental Difference Equation SIR Models

Mauricio Santillana Ashleigh Tuite Tahmina Nasserie
 Paul Fine David Champredon Leonid Chindelevitch
 Jonathan Dushoff David Fisman

March 4, 2016

Abstract

Mathematical models are often regarded as recent innovations in the description and analysis of infectious disease outbreaks and epidemics, but simple models have been in use for projection of epidemic trajectories for more than a century. We recently described a single equation model (the incidence decay with exponential adjustment, or IDEA, model) that can be used for short term forecasting. In the mid-19th century, Dr. William Farr developed a single equation approach (“Farr’s law”) for epidemic forecasting. We show here that the two models are in fact identical, and can be expressed in terms of one another, and also in terms of a susceptible-infectious-removed (SIR) compartmental model with improving control. This demonstrates that the concept of the reproduction number (\mathcal{R}_0) is implicit to Farr’s (pre-microbial era) work, and also suggests that control of epidemics, whether via behavior change or intervention, is as integral to the natural history of epidemics as is the dynamics of disease transmission.

1 Introduction

The control of communicable diseases is an endeavor that has witnessed remarkable successes over the past century; diseases that previously caused large scale mortality have been eradicated [1, 2], locally eliminated [3], or have been markedly reduced in incidence globally as a result of vaccination, antimicrobial therapy, water and sewage treatment, and advances in food safety [4–6]. Nonetheless, the threat of communicable diseases persists; emerging infectious diseases continue to be identified, often in association with changes in human and animal mobility, agricultural practices, environmental degradation, and misuse of antimicrobial therapy [7–9]. Recent outbreaks or epidemics associated with MERS coronavirus [10], influenza A(H7N9) [11], and the West African emergence of the Zaire strain of Ebola virus [12], have challenged epidemiologists as the natural history, modes of transmission, and/or means of control of

these diseases have not been well understood during initial periods of emergence.

When novel infectious diseases emerge or familiar diseases resurge, mathematical models can serve as useful tools for the synthesis of available data, management of uncertainty, and projection of likely epidemic trajectories [13]. While it may be challenging to parameterize detailed mechanistic mathematical models when there is little information on mechanisms of transmission, baseline immunity in a community, the nature of the infecting pathogen, et cetera, a number of descriptive approaches exist which may permit fitting, and forecasting, of an epidemic curve. One single equation approach that has been applied to emerging infections is the Richards model, which treats cumulative infections as a logistic growth process [14]. However, the concept of modeling an epidemic curve as a simple function, without reference to mechanisms of transmission, is in fact much older, and may originate in the work of the English polymath Dr. William Farr (1807-1883), who rose from humble beginnings to become a physician, mathematician, hygienist and protege of Lancet founder Dr. Thomas Wakley [15–17]. Dr. Farr spent almost 40 years at the General Register Office of the United Kingdom, and the esteem in which he was held is apparent in the "letters" he published annually as appendices to the reports of the Registrar General, in which he supplemented the dry statistical reports with thoughtful and creative musings on topics as wide-ranging as the relationships between occupation and disease, suicide and mortality in the mentally ill, population density and mortality, and as above the "laws" governing epidemics [18].

Using what was subsequently dubbed "Farr's law" [17, 19], Farr used the ratio of case counts in successive time periods to successfully forecast the size and duration of epidemics of smallpox and rinderpest [16, 17] (Figure 1). Farr demonstrated in a "letter" of 1837 that the decline of a national smallpox epidemic was exponential rather than linear [18]. He used related methods over 25 years later, to refute the assertion by a British parliamentarian that a rinderpest epidemic in British cattle would continue to grow exponentially over time [16, 17]. Again, Farr demonstrated that although case numbers were increasing markedly, the rate of increase was, in fact, decelerating; he asserted that the epidemic would be large but would end in the coming months, a projection which proved to be correct.

Farr was famously vague about the mathematical underpinnings of his projections, though key to his observation was the observation that case counts would decrease by a constant log-quantity following initial exponential epidemic growth [16]. It fell to other contemporaries [26] and later epidemiologists (most notably Dr. John Brownlee) to formalize "Farr's law" [15, 27, 28]. (It should be noted that the term "Farr's law" is ambiguous. Farr himself referred to a "law" in his letter on rinderpest [16], but the term has also been used by others to describe Farr's observations on the relation between population density and death [20], and to his description of the relationship between cholera mortality and altitude [21]. In his elaboration of the "law", Brownlee referred to it as "Farr's theory of epidemics" [16]).

We recently proposed a descriptive approach to the initial estimation of the basic reproduction number (\mathcal{R}_0) of an emerging or re-emerging pathogen, which also provides information on the rate at which the process is being controlled, as well as reasonable short-term projections of incidence. This two-parameter model, which we have referred to as the “Incidence Decay with Exponential Adjustment” (IDEA) model, offers advantages of simplicity, explicit linkage to theory of epidemic growth, and also acknowledges the fact that epidemics and outbreaks do not peak and end simply due to depletion of susceptibles, but because of a complex constellation of public health actions and behavioral changes that may modify the course of an epidemic and reduce the effective reproductive number $\mathcal{R}_e(t)$ of an outbreak [22]. One of us (PF) had previously written about Farr’s rule and its importance in the development of epidemic theory [15], and noted the conceptual similarity between IDEA and Farr’s rule. Upon exploration of these two approaches we realized that they are, notwithstanding having been formulated some 160 years apart, and being based on very different theoretical constructs, identical. Here we demonstrate the equivalence of these methods for epidemic modeling and forecasting, and explore the potential advantages of using these methodologies to complement other approaches. We explore numerical examples derived from current emerging infectious diseases, including the West African Ebola virus outbreak. We show that these methodologies work fairly well in a number of cases, and discuss what this fact implies about difficult-to-measure components of transmission and control parameters.

2 Methods

2.1 Farr’s Rule

The empirical relationship between observed cases of infected individuals, in sequential time intervals, during an epidemic outbreak according to **Farr’s rule** is given by

$$\frac{\frac{I(t+3)}{I(t+2)}}{\frac{I(t+2)}{I(t+1)}} = K \quad (1)$$

where $I(t)$ represents the number of observed newly infected individuals at time t , and K is a constant. For values of $K < 1$, the rate of change in the observed cases (acceleration) decreases as time evolves, and the family of curves $I(t)$ that satisfy equation (1) correspond to familiar bell-shaped epidemiological curves.

The rule was investigated and elaborated upon by the epidemiologist John Brownlee who noted that $I(t)$ under this rule would correspond to the function [15]

$$\exp(-At^2 + Bt + C) \quad (2)$$

where A , B , and C are constants. Of note, Brownlee’s formulation identifies a process where cases increase as a first order process, but decrease as a second order process, as is the case with the IDEA model.

2.2 IDEA Model

In its basic form the IDEA model holds

$$I(t) = \left(\frac{\mathcal{R}_0}{(1+d)^t} \right)^t \quad (3)$$

where $I(t)$ is the incident infections in generation t of an outbreak (thus $t = \{1, 2, 3, \dots\} \in \mathbb{N}$). This is in contradistinction to equation (1) above where t is an arbitrary, but constant, time interval (not necessarily equal to a generation interval). The parameter \mathcal{R}_0 is the basic reproduction number as usually defined; that is, the number of secondary cases created by a primary case in a totally susceptible population and in the absence of intervention. The parameter d which we have referred to as a “control parameter” defines the rate at which transmission declines over the course of an epidemic. The empirical underpinnings of d are not yet well defined, but based on current understanding of disease dynamics could represent public health interventions, population adaptation or behavior change, improved availability of personal protective items or effect of drugs to treat infection, or reductions in population susceptibility as a result of infection or vaccination. As noted above, by fitting the model to data we have previously obtained reasonable estimates of \mathcal{R}_0 early in the course of epidemics, and have also been able to produce plausible near-term projections of future case counts.

2.3 Difference Equation Susceptible-Infectious-Removed Model

In an earlier publication we commented on the almost identical projections generated by IDEA and a compartmental difference equation (Susceptible-Infectious-Removed, or SIR) [22], when \mathcal{R}_0 is small and when there is exponential reduction in risk over time. Indeed, the effect we identified can be generalized to any situation where depletion of susceptibles due to infection is small relative to the total population size, and not only when \mathcal{R}_0 is small. We used a “damped” version of the standard SIR model whose formulation in generation interval time scale is given by:

$$S_{t+1} = S_t - \mathcal{R}_e(t)I_t \quad (4)$$

$$I_{t+1} = \mathcal{R}_e(t)I_t \quad (5)$$

with S_t the number of susceptible individuals at time t and $\mathcal{R}_e(t)$ the effective reproductive number at time t defined by:

$$\mathcal{R}_e(t) = \mathcal{R}_0 \frac{S_t}{N} \rho^t \quad (6)$$

The “dampening” parameter ρ represents the relative risk of infection in each generation of the epidemic, compared to risk seen in the last generation (i.e., if there were no improvement in control in a given generation compared to the last, most recent generation). If an outbreak is small relative to the size of the total population (as would be true if \mathcal{R}_0 were modest and control achieved relatively quickly) $S(t)/N$ will be approximately 1 throughout the outbreak and the expression can be rewritten as:

$$I(t) = I(0) \prod_{s < t} (\rho^s \mathcal{R}_0) \quad (7)$$

meaning that all reduction in $\mathcal{R}_e(t)$ is due to control rather than depletion of susceptibles. \mathcal{R}_0 is a constant, and the sum of exponents of ρ is simply $t(t+1)/2$. We can assume (as we do with IDEA) that the outbreak began with the introduction of a single case such that $I(0) = 1$. Now:

$$I(t) = \rho^{t(t+1)/2} \mathcal{R}_0^t \quad (8)$$

2.4 Numerical Examples using Recent Outbreak Data

We have used IDEA to explore the nature of epidemic growth during the recent West African Ebola epidemic [31, 32], the 2014 MERS coronavirus outbreak in Saudi Arabia [33], and more recently Chikungunya virus invasion in the Western hemisphere [34]. As publicly available data have taken the form of cumulative incidence curves (with absent dates of onset) we fit IDEA to cumulative curves, but it is possible to estimate “pseudo-incidence” by taking the interval to interval difference cumulative incidence over time.

We fit IDEA to the incidence time series and calculated Farr’s K for sequential generation tetrads, and converted K values to the d parameter in IDEA using the relation $K = \frac{1}{(1+d)^4}$ described below. Of necessity, we excluded K estimates derived from sequences of intervals containing negative incidence estimates (i.e., those where reported cumulative incidence declined). Data sources used for these analyses are available at [http://figshare.com/authors/Tahmina_Nasserie/686527\(Chikungunya\)](http://figshare.com/authors/Tahmina_Nasserie/686527(Chikungunya)) and <https://github.com/cmrrivers/ebola> (Ebola).

3 Results

3.1 Equivalence of IDEA and Farr’s Rule

An incidence curve described by IDEA naturally satisfies Farr’s rule. Indeed, substituting (3) into (1) gives

$$K = \frac{1}{(1+d)^4} \quad (9)$$

It is interesting to note that the value of \mathcal{R}_0 in the IDEA model is irrelevant in the proof (see Appendix 5.1).

Moreover, by expressing the IDEA model as

$$I(t) = \left(\frac{\mathcal{R}_0}{(1+d)^t} \right)^t = \exp \left(t \log \frac{\mathcal{R}_0}{(1+d)^t} \right) = \exp(-t^2 \log(1+d) + t \log \mathcal{R}_0) \quad (10)$$

we see that we recover John Brownlee's Gaussian curve (2) with $A = \log(1+d)$, $B = \log \mathcal{R}_0$ and $C = 0$.

3.2 K as an Odds Ratio

As K represents a ratio of ratios, it can be conceptualized as equivalent in form to an odds ratio. Effectively, $\frac{I_t}{I_{t+1}}$ is the odds of a case occurring in an initial as opposed to a subsequent generation while K then becomes interpretable as an odds ratio, though it is unclear whether this odds ratio has an intuitive meaning. Nonetheless, this form is important as it suggests that the asymptotic variance of $\log(K)$ can be estimated as $(\frac{1}{I_t} + \frac{1}{I_{t+1}} + \frac{1}{I_{t+2}} + \frac{1}{I_{t+3}})$ [35]. Estimation of variance of $\log(K)$ makes it possible to estimate confidence limits for K , and thus for d . Furthermore, given that several K estimates can be generated from a given epidemic curve, estimation of variance should permit the use of Mantel-Haenszel methods [37] to generate summary estimates of K over the course of an epidemic, or to use meta-regressive methods to evaluate K for trends [36]. A potential pitfall here is the non-independence of serial estimates of K due to overlap in incidence values used in adjacent estimates of K .

3.3 Relation of SIR Model to Farr's K and IDEA

When the depletion of susceptible is negligible compared to the total population size (that is, a small outbreak), we can actually express IDEA's basic reproductive number $\mathcal{R}_{0,IDEA}$ and its control factor d as a function of the basic reproductive number $\mathcal{R}_{0,SIR}$ and the control factor ρ of the damped SIR model described in (4-6). The relationship between the parameters of these two models is (details in Appendix 5.2)

$$\mathcal{R}_{0,IDEA} \simeq \frac{\mathcal{R}_{0,SIR}}{\sqrt{\rho}} \quad (11)$$

$$d \simeq \frac{1}{\sqrt{\rho}} - 1 \quad (12)$$

Substituting (12) in equation (9) we can link Farr's rule with the damped SIR model:

$$K \simeq \rho^2 \quad (13)$$

3.4 Numerical Simulations

In this section, we aim to test numerically the validity of approximations (11) and (12). In particular, given a damped SIR model, we explore the parameter

space $\mathcal{R}_{0,SIR}$ and ρ for which these approximations hold. Note that the link between IDEA and Farr’s rule given by equation (9) is not an approximation, but a genuine equality (subject to the time step used in Farr’s rule being equivalent to one generation), so there is no need to test it numerically. To measure the performance of the approximation, we consider the distance between the simulated incidence time series. Let N be the number of generations simulated and $I_{SIR}(k)$ (resp. $I_{IDEA}(k)$) the incidence from the SIR (resp IDEA) model at the k^{th} generation, we define their distance by

$$\delta = \sqrt{\sum_{k=1}^N (I_{SIR}(k) - I_{IDEA}(k))^2} \quad (14)$$

The tests are implemented in the software R [39] and code is available in electronic supplement files. Figure 3 shows the values of the distance δ for different values of $\mathcal{R}_{0,SIR}$ and ρ . We see that for a combination of $\mathcal{R}_{0,SIR}$ and ρ such that the depletion of susceptible is not too large, the approximation is very good. But when the values of $\mathcal{R}_{0,SIR}$ and ρ generate a depletion of susceptible individuals that is no longer negligible (white area in Figure 3), then the incidence curve from the IDEA model diverges from that generated by the damped SIR model.

3.5 Application: Ebola and Chikungunya

Considering real epidemic data from Ebola and Chikungunya, it can be seen that the interval to interval variability in K was substantial, likely reflecting variability in reporting (Figure 5 and Figure 6). Simple arithmetic means of K over time were also unstable due to skewing by values substantially greater than 1. However, when we estimated the geometric mean of K over time we found that the resultant d estimate approximated that derived through fitting IDEA.

Furthermore, we noted that in the Chikungunya time series there was a large perturbation in best-fit values of d occurring in October 2014, corresponding with an apparent multi-wave epidemic. We have previously noted that this abrupt change the generation-to-generation best fit value of d corresponds with the occurrence of multiwave epidemics when IDEA is fit to simulated data [22]; using Farr’s approach, the onset of a possible new Chikungunya wave seems to correspond with an abrupt increase in K to a value far greater than one (Figure 7). The utility of large values of K as a signal of an incipient epidemic wave warrants further investigation.

4 Discussion

Although the real-time application of mathematical modeling to understanding and control of outbreaks is often perceived as representing a recent development in infectious disease epidemiology [23], disease modeling has deeper historical

roots, including work by Bernoulli on smallpox in the 18th century [24]; work by Ross on malaria transmission [25], and as mentioned above, Farr’s work on the growth and cessation of epidemics [15, 16, 27]. We had published a simple, phenomenological approach to the description and projection of outbreaks and epidemics [22] which we had initially regarded as a novel formulation rooted in the concept of the basic reproduction number \mathcal{R}_0 .

In that work, we demonstrated concordance with projections derived using a 3-compartment difference equation model (damped SIR model). We have subsequently realized, and demonstrate above, that our approach simply represented a restatement of Farr’s work, albeit in a manner that is tied to the concept of \mathcal{R}_0 . According to Brownlee, Farr promised to describe the derivation of his model in greater detail in future reports, but never did so [16], and much of the mathematical elaboration of Farr’s work was in fact done by Brownlee after Farr’s death [27]. Nonetheless, Brownlee notes that to Farr, the predictive accuracy of his approach reflected three characteristics of epidemics, according to Farr’s (pre-microbial) understanding: (i) diminution in the number of susceptibles over time due to recovery from infection (“immunity”, though to use this term in application to Farr is an anachronism); (ii) diminished population density due to death from infection; and (iii) diminishing pathogenicity of the disease with each passing generation of infection as a result of (to quote Farr) “[loss of] part of the force of infection in every body through which they pass...the matter...diminishes in strength at every transmission by inoculation” [16]. The first two characteristics are not incompatible with the current understanding of epidemic dynamics, whereas the third is not (though it does anticipate more modern ideas around evolution of virulence and disease ecology [29, 30]).

However, as this model is phenomenological, rather than mechanistic in nature, the putative epidemiological mechanisms underlying model performance are not of immediate importance. Indeed, while the simplicity of this approach may be regarded as a limitation, the simplicity of the form, and its implicit incorporation of biological, social, medical, and behavioral drivers of control into a single parameter estimated via fitting, may be a strength, especially given that such control factors as behavior change due to fear may be difficult or impossible to measure in real time.

When we have applied IDEA to current day outbreaks and epidemics, we have remained agnostic about the factors that cause second order deceleration of epidemic growth. Referring to first principles, the components of a reproduction number are duration of infectiousness, contact rate, and probability of transmission conditional on contact, as well as susceptibility in a population [38]. We presume that public health interventions, population behavior change (as a result of education or rumours, prudence or fear), and the occurrence of silent infections with immunity could all contribute to deceleration of epidemic spread, even when the effective reproductive number is expected to be greater than 1 due to widespread susceptibility in the population. Furthermore, we note that Farr’s original “time step” appears to have been arbitrary (reflecting the form of data available to him: weekly for rinderpest, quarterly for smallpox), whereas we have used generations as time steps in our more recent applications of Farr’s “law”, and in IDEA and SIR models. This potential discrepancy between Farr’s initial efforts and our more recent efforts warrants further exploration.

4.1 Summary

In demonstrating that the IDEA model and Farr’s model are mathematically identical (and can be virtually identical to an SIR model with a small \mathcal{R}_0 , abundant susceptibility in the population, and exponential improvement in control) we demonstrate that recognizing the underlying mechanism of epidemic control may be unimportant for generating reasonable forecasts of epidemics with control, or recognizing when their fundamental dynamics have changed. Our contribution in the current work is to show that Farr’s law, while derived in the pre-microbial era, can be reformulated in terms of the concept of the basic reproductive number, combined with exponential increase in control via whatever mechanism. We observe, unexpectedly, that Farr’s K can be expressed as a function of the IDEA d parameter alone, independent of \mathcal{R}_0 , implying that epidemic trajectory is (and has historically been) more a function of control efforts and changing behavior than of the fundamental characteristics of a given infectious disease. Whether or not the ratio K can have stand-alone value as a tool to identify unexpected shifts in epidemic trajectory (e.g., the two wave epidemic of Chikungunya referred to in Figure 7 above) will be the subject of future work.

5 Appendix

5.1 Details on IDEA and Farr's law relationship

[DC: Depending of the journal this manuscript is submitted to, I would either (if slightly technical journal) not include this proof at all or (non-technical journal) do not include case $t = 1$]

Assume that the parameters \mathcal{R}_0 and $d \geq 0$, in the IDEA model, describe accurately the observed number of infected cases as a function of (serial) time, in an ideal outbreak. Then we can substitute $I(t)$ from equation (3) into equation (1) to see if the sequential number of infected individuals, as predicted by the IDEA model, satisfy Farr's law.

For clarity, choose $t = 1$, thus equation (1) becomes

$$\begin{aligned} \frac{I(4)}{I(3)} &= \frac{\left(\frac{\mathcal{R}_0}{(1+d)^4}\right)^4}{\left(\frac{\mathcal{R}_0}{(1+d)^3}\right)^3} = \frac{\left(\frac{\mathcal{R}_0^4}{\mathcal{R}_0^3}\right) \left(\frac{((1+d)^3)^3}{((1+d)^4)^4} f\right)}{\left(\frac{\mathcal{R}_0^2}{\mathcal{R}_0}\right) \left(\frac{(1+d)}{((1+d)^2)^2}\right)} \\ &= \frac{\mathcal{R}_0 \left(\frac{(1+d)^9}{(1+d)^{16}}\right)}{\mathcal{R}_0 \left(\frac{(1+d)}{(1+d)^4}\right)} = \frac{(1+d)^{-7}}{(1+d)^{-3}} = \frac{1}{(1+d)^4} \end{aligned}$$

By identifying the constant $K = 1/(1+d)^4$, the IDEA model satisfies Farr's law for $t = 1$. In general, for any sequential (integer) time intervals $t, t+1, t+2, t+3$ one can generalize the above result as follows:

$$\begin{aligned} \frac{I(t+3)}{I(t+2)} &= \frac{\left(\frac{\mathcal{R}_0}{(1+d)^{t+3}}\right)^{t+3}}{\left(\frac{\mathcal{R}_0}{(1+d)^{t+2}}\right)^{t+2}} = \frac{\left(\frac{\mathcal{R}_0^{t+3}}{\mathcal{R}_0^{t+2}}\right) \left(\frac{((1+d)^{t+2})^{t+2}}{((1+d)^{t+3})^{t+3}}\right)}{\left(\frac{\mathcal{R}_0^{t+1}}{\mathcal{R}_0^t}\right) \left(\frac{((1+d)^t)^t}{((1+d)^{t+1})^{t+1}}\right)} \\ &= \frac{\mathcal{R}_0 \left(\frac{(1+d)^{(t+2)^2}}{(1+d)^{(t+3)^2}}\right)}{\mathcal{R}_0 \left(\frac{(1+d)^{t^2}}{(1+d)^{(t+1)^2}}\right)} = \frac{(1+d)^{(t+2)^2 - (t+3)^2}}{(1+d)^{t^2 - (t+1)^2}} \\ &= \frac{(1+d)^{-2t-5}}{(1+d)^{-2t-1}} = \frac{1}{(1+d)^4} \end{aligned}$$

5.2 Details on IDEA and SIR relationship

Incidence or the SIR model (Equation 4) can be written as

$$\begin{aligned}
I_t^{SIR} &= RE(t)I_{t-1} \\
&= I_0 \prod_{h=0}^{t-1} RE(s) \\
&= I_0 \prod_{h=0}^{t-1} R_{0,SIR} S_h/N \rho^h \\
&= R_{0,SIR}^t \rho^{t(t-1)/2} \prod_{h=0}^{t-1} S_h/N
\end{aligned}$$

If the epidemic size is small compared to the size of the whole population, then it can be assumed that

$$\prod_{h=0}^{t-1} \left(\frac{S_h}{N} \right) \simeq 1$$

In that case,

$$I_t^{SIR} \simeq \mathcal{R}_{0,SIR}^t \rho^{t(t-1)/2} = \frac{(\mathcal{R}_{0,SIR}/\sqrt{\rho})^t}{(1/\sqrt{\rho})^{t^2}} \quad (15)$$

Incidence in the IDEA framework is simply

$$I_t^{IDEA} = \left(\frac{\mathcal{R}_{0,IDEA}}{(1+d)^t} \right)^t = \frac{\mathcal{R}_{0,IDEA}^t}{(1+d)^{t^2}} \quad (16)$$

Finally, both models have the at time t if $I_t^{SIR} = I_t^{IDEA}$. A sufficient condition for that equality to hold is when both numerator and denominator of equations (15) and (16) are equal, that is

$$\mathcal{R}_{0,IDEA} \simeq \frac{\mathcal{R}_{0,SIR}}{\sqrt{\rho}}$$

$$d \simeq \frac{1}{\sqrt{\rho}} - 1$$

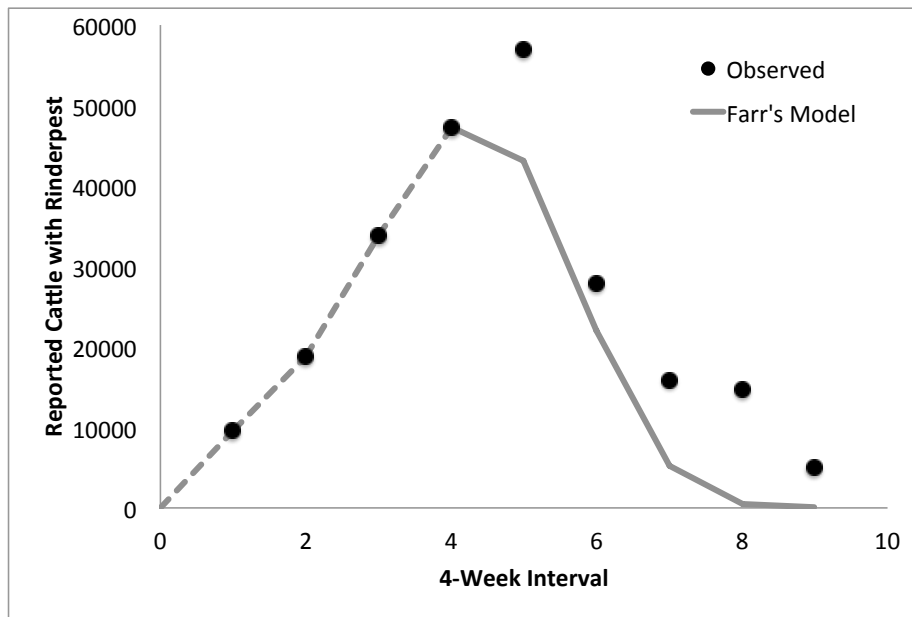


Figure 1: A recreation of Farr's possible approach to the English rinderpest epidemic of 1865-1866, based on data presented in [16]. Black dots = empirical data, whereas gray curve represent Farr's "model projections". Dashed component of curve represents fitted component of "model", while solid gray curve represents Farr's projection. Initial data point is November 4, 1865, with subsequent points presented at 4 week intervals to June 16, 1866.

6 Figures

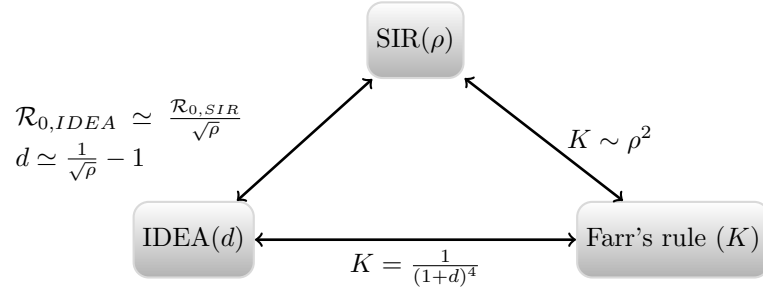


Figure 2: Relationship between the three models. Parameter ρ represent the relative risk of infection in each generation for the SIR model. The control parameter d is associated with the IDEA model and K is Farr's ratio.

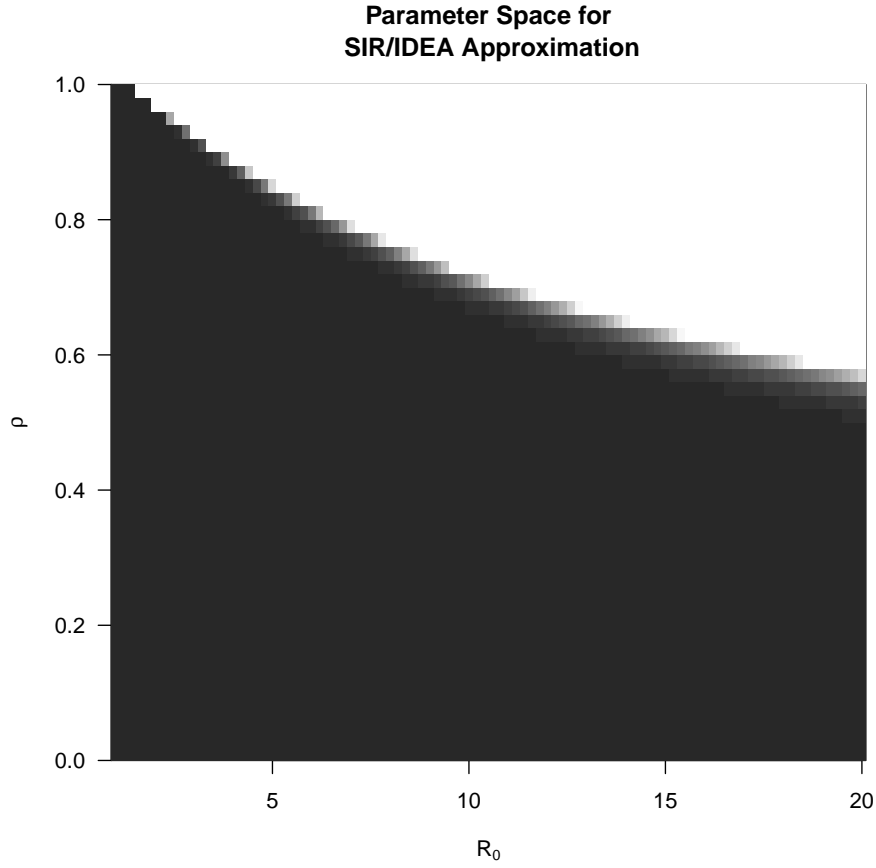


Figure 3: A heat map plotting values for $\mathcal{R}_{0,SIR}$ and ρ where the damped SIR model can be approximated by a IDEA model using equations (11) and (12). Darker areas indicate a good match (measured as the sum of squared differences) between the simulated incidence time series; lighter areas represent combinations of values for which incidence time series for SIR and IDEA diverge.

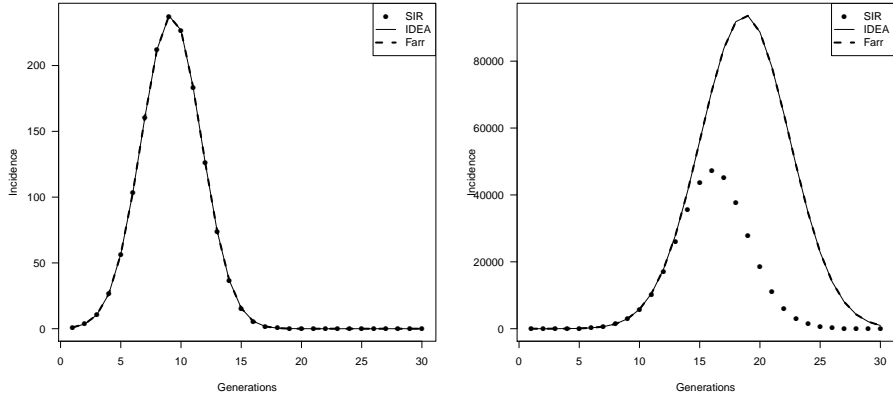


Figure 4: Evaluation of IDEA model fit to simulated data derived from the damped SIR model. Left panel illustrates a scenario where approximation is very good, ($\mathcal{R}_{0,SIR}=3.5$, $\rho=0.85$), corresponding to a combination of values found in the dark area in Figure 3. The right sided panel uses a combination of values (high $\mathcal{R}_{0,SIR}$ and/or low ρ) where susceptible depletion cannot be ignored (i.e., corresponding to the white area in Figure 3). It can be seen that IDEA and the damped SIR models diverge when susceptibles are rapidly depleted.

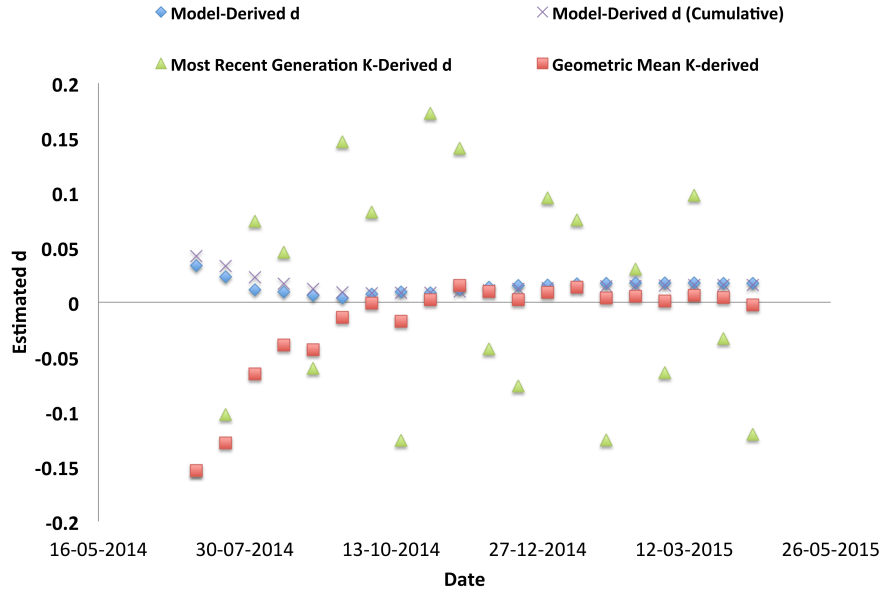


Figure 5: The graph plots estimates of IDEA d parameter against time during the recent West African Ebola outbreak. Approximate date of the last generation incorporated into estimates is plotted on the X-axis; estimated d is plotted on the Y-axis. d estimates were either derived via IDEA model fitting to “incident” cases (blue diamonds) or cumulative incidence (crosses), or derived by estimating Farr’s K and transforming resultant estimates using the relation described by equation (9). When K is estimated using 4-generation series (green diamonds), resultant d estimates are volatile and bear little resemblance to d estimates derived through fitting IDEA. However, estimates of K derived as geometric means of all available K values (red squares) provide a more reasonable approximation of d .

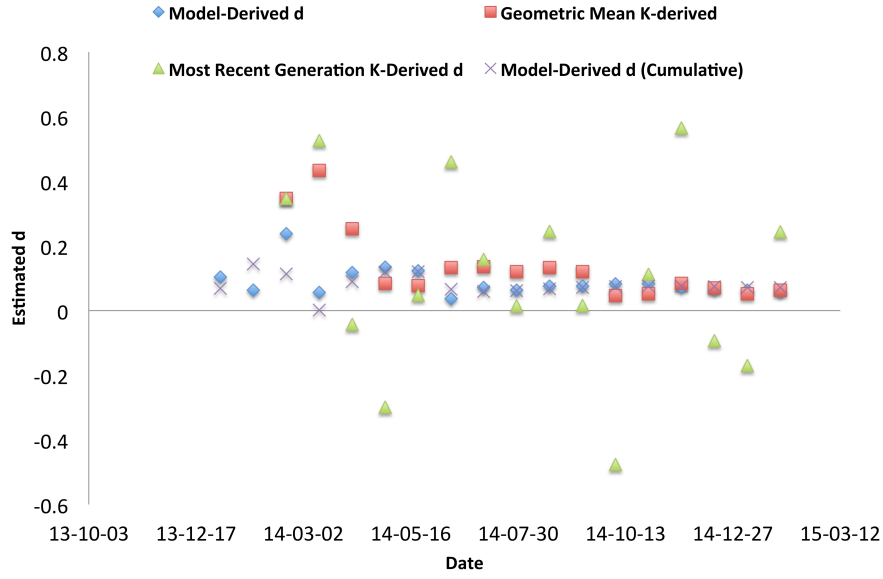


Figure 6: As with Figure 5, this figure shows estimates of d , derived directly by model fitting or by transforming estimates of Farr’s K , for the emerging Western Hemisphere Chikungunya epidemic in 2014-2015. As in Figure 5, d estimates were either derived via IDEA model fitting to “incident” cases (blue diamonds) or cumulative incidence (crosses), or derived by estimating Farr’s K and transforming resultant estimates. As in Figure 5, volatile estimates of K were derived using 4-generation series (green diamonds), but estimates of K derived as geometric means of all available K values (red squares) provided a reasonable approximation of d .

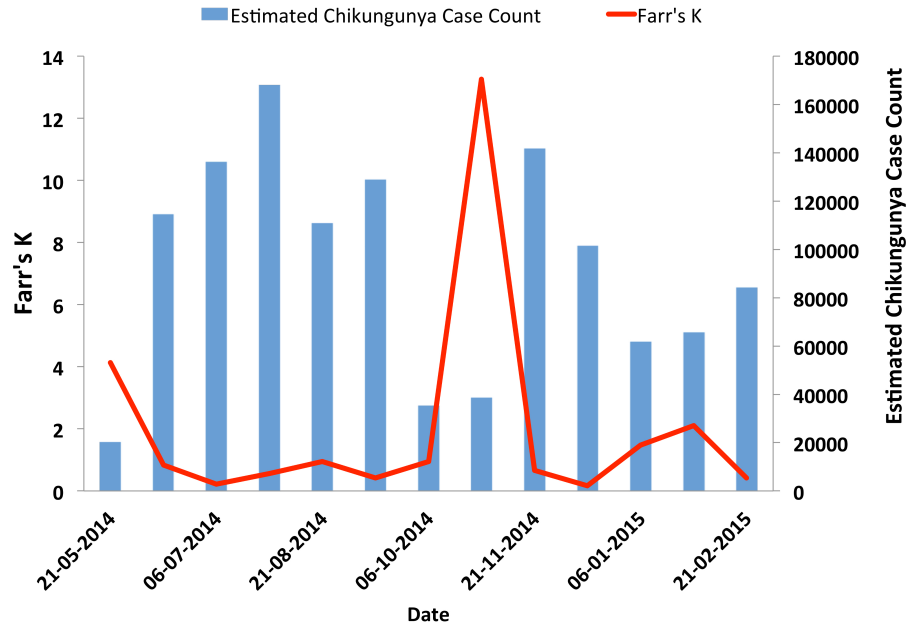


Figure 7: A possible application for raw estimates of Farr's K emerged in analysis of data from the 2014-2015 Western Hemisphere Chikungunya outbreak; here it appears that a multi-wave epidemic is signaled by a sudden surge in K to a value > 1 (red line), indicating that there is renewed exponential growth in cases (blue bars), rather than exponential decline. X-axis, date of most recent generation; left Y-axis, Farr's K ; right Y-axis, estimated per-generation Chikungunya case count and transforming resultant estimates. As in Figure 1, volatile estimates of K were derived using 4-generation series (green diamonds), but estimates of K derived as geometric means of all available K values (red squares) provided a reasonable approximation of d .

References

- [1] Hinman A. Eradication of vaccine-preventable diseases. *Annual Review of Public Health* 1999; 20: 211-229.
- [2] Roeder P, Mariner J, Kock R. Rinderpest: the veterinary perspective on eradication. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences* 2013; 368: 20120139.
- [3] Papania MJ, *et al.* Elimination of endemic measles, rubella, and congenital rubella syndrome from the Western Hemisphere: the US experience. *Journal of the American Medical Association Pediatrics* 2014; 168: 148-155.
- [4] Armstrong GL, Conn LA, Pinner RW. Trends in infectious disease mortality in the United States during the 20th century. *Journal of the American Medical Association* 1999; 281: 61-66.
- [5] Murray CJ, *et al.* Global, regional, and national incidence and mortality for HIV, tuberculosis, and malaria during 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2014; 384: 1005-1070.
- [6] Liu L, *et al.* Global, regional, and national causes of child mortality in 2000-13, with projections to inform post-2015 priorities: an updated systematic analysis. *Lancet* 2015; 385: 430-440.
- [7] Kuehn BM. Human, animal, ecosystem health all key to curbing emerging infectious diseases. *Journal of the American Medical Association* 2010; 303: 117-118, 124.
- [8] Keesing F, *et al.* Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature* 2010; 468: 647-652.
- [9] Jones KE, *et al.* Global trends in emerging infectious diseases. *Nature* 2008; 451: 990-993.
- [10] Azhar EI, *et al.* Evidence for camel-to-human transmission of MERS coronavirus. *New England Journal of Medicine* 2014; 370: 2499-2505.
- [11] Cowling BJ, *et al.* Comparative epidemiology of human infections with avian influenza A H7N9 and H5N1 viruses in China: a population-based study of laboratory-confirmed cases. *Lancet*— 2013; 382: 129-137.
- [12] Baize S, Pannetier D, Oestereich L, *et al.* Emergence of Zaire Ebola virus disease in Guinea. *New England Journal of Medicine* 2014; 371: 1418-1425.
- [13] Fisman D. Modelling an influenza pandemic: A guide for the perplexed. *Canadian Medical Association Journal* 2009; 181: 171-173.
- [14] Hsieh YH, Chen CW. Turning points, reproduction number, and impact of climatological events for multi-wave dengue outbreaks. *Tropical Medicine & International Health* 2009; 14: 628-638.
- [15] Fine PEM. John Brownlee and the measurement of infectiousness: An historical study in epidemic theory. *Journal of the Royal Statistical Society. Series A* 1979; 142: 347-362.

- [16] Brownlee J. Historical note on Farr's theory of the epidemic. *British Medical Journal* 1915; 2: 250-252.
- [17] Greenwood M. William Farr. *Lancet* 1933; 5724: 1047-1052.
- [18] Farr, W. Appendix to the second annual report of the registrar-general. London, UK. General Register Office, 1837. *Online Historical Population Reports*. Available via the Internet at <http://histpop.org/ohpr/servlet/PageBrowser2?ResourceType=Registrar>
- [19] Langmuir AD, Bregman DJ. Farr's law applied to AIDS projections. *Journal of the American Medical Association* 1990; 263: 1522-1525.
- [20] Brownlee J. Studies in the meaning and relationships of birth and death rates II: density of population and death rate (Farr's law). *The Journal of Hygiene* 1915; 15: 11-6.
- [21] Lilienfeld DE. Celebration: William Farr (1807–1883) - An appreciation on the 200th anniversary of his birth. *International Journal of Epidemiology* 2007; 36: 985-987.
- [22] Fisman DN, *et al.* An IDEA for short term outbreak projection: nearcasting using the basic reproduction number. *Public Library of Science One* 2013; 8: e83622.
- [23] Heesterbeek H, *et al.* Modeling infectious disease dynamics in the complex landscape of global health. *Science* 2015; 347: aaa4339.
- [24] Greenwood M. Medical Statistics from Graunt to Farr. *Biometrika* 1941; 32: 101-127.
- [25] Smith DL, *et al.* Ross, Macdonald, and a theory for the dynamics and control of mosquito-transmitted pathogens. *Public Library of Science Pathogens* 2012; 8: e1002588.
- [26] Evans G.H. Some arithmetical considerations on the progress of epidemics. *Transactions of the Epidemiological Society of London*, Sessions 1866-1876, volume 3. 551-555. Hardwicke and Bogue, London, 1876.
- [27] Brownlee J. On the curve of the epidemic. *British Medical Journal* 1915; 2: 799-800.
- [28] Serfling R.E. Historical review of epidemic theory. *Human Biology* 1952 24(3):145-66.
- [29] Lipsitch ML, Moxon ER. Virulence and transmissibility of pathogens: what is the relationship? *Trends in Microbiology* 1997; 5: 31-37.
- [30] Ewald P. Evolution of virulence. *Infectious Disease Clinics of North America* 2004; 18: 1-15.
- [31] Fisman DN, Khoo E, Tuite AR. Early epidemic dynamics of the west african 2014 ebola outbreak: estimates derived with a simple two-parameter model. *Public Library of Science Currents*. Published online: 8 September 2014. doi: 10.1371/currents.outbreaks.89c0d3783f36958d96ebbae97348d571.

- [32] Fisman DN, Tuite A. Projected impact of vaccination timing and dose availability on the course of the 2014 West African Ebola epidemic. *Public Library of Science Currents*. Published online: 21 November 2014. doi: 10.1371/currents.outbreaks.06e00d0546ad426fed83ff24a1d4c4cc.
- [33] Majumder MS, *et al.* Estimation of MERS-Coronavirus reproductive number and case fatality rate for the spring 2014 Saudi Arabia outbreak: Insights from publicly available data. *Public Library of Science Currents* Published online: 18 December 2014. doi: 10.1371/currents.outbreaks.98d2f8f3382d84f390736cd5f5fe133c.
- [34] Nasserie T, Fisman DN, Tuite AR. Using the Incidence Decay with Exponential Adjustment Model to understand Chikungunya epidemic growth in the Americas [conference presentation]. Canadian Society for Epidemiology and Biostatistics 2015 Conference. Mississauga, Ontario, Canada. June 1-4, 2015.
- [35] Kelsey JL, *et al.* Case control studies II: further design considerations and analysis. In: *Methods in Observational Epidemiology*, 2nd edn. New York: Oxford University Press, 1996. pp. 214-243.
- [36] Egger M, Smith GD, Altman DG. Going beyond the grand mean: subgroup analysis in meta-analysis of randomised trials. In: *Systematic Reviews in Healthcare: Meta-Analysis in Context*. London: BMJ Books, 2001. pp. 143-156.
- [37] Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute* 1959; 22: 719-748.
- [38] Vynnycky E, White RG. The short term dynamics of infections. In: *An Introduction to Infectious Disease Modelling*. New York: Oxford University Press, 2010.
- [39] The R Project for Statistical Computing. Available via the Internet at <https://www.r-project.org/>. Last accessed November 25, 2015.